

# Memo

---

To: Bob North, Bob Woodward

From: Xiaoping Yang, Keith McLaughlin, and Istvan Bondar

Date: July 24, 2001

Subject: 90% error ellipse coverage testing

CC: Jack Murphy

---

It is often claimed that the modeling error estimates currently in use in the IDC location process need improvement (Figure 1), based on selected comparisons of outlying GSETT-3 locations with local network locations. However, no evaluation has been done using a large set of accurate ground truth (GT) events to evaluate coverage performance.

A large set of GT0-GT10 events was recently collected for the Mediterranean, North Africa, Middle East, and Western Eurasia by the Group-2 Location Calibration Consortium. These GT origins and arrivals were collected to test and validate 3D models. Arrivals include both regional and teleseismic P and S phases.

A by-product of the Group-2 testing procedures was a test of the existing IDC model errors. Over 600 GT0-GT10 events were used in the testing, with at least three defining Pn and Sn phases (Table 1; Figure 2). We relocated these events using the IDC IASPEI91 travel-time tables by running the EvLoc location program, consistent with standard IDC practices. Table 2 shows the 90% ellipse coverage for multiple tests using this data set. For all tests the 90% ellipse coverages are over 90%. The tests validate the existing combined and separate regional and teleseismic model errors with regard to providing honest 90% ellipses.

The 90% error ellipse coverage is scaled by the GT uncertainty. When using all regional and teleseismic phases, the 90% coverage was 93%. This means 580 out of 625 GT origins located within the error ellipse, very close to the expected 563 origins. It indicates that the modeling error estimates in use are reasonable, and the relative uncertainties of regionals and teleseismics appear to be correct.

Figure 3 shows a cumulative plot of the 90% error ellipse statistic coverage for 625 events relocated using the IDC IASPEI91 travel-time tables, compared to the theoretical  $\chi^2$  distribution with 2 degrees of freedom. The Kolmogorov-Smirnov test value is 0.28 indicating fairly high risk for the assumption that the two distributions are similar. In fact, median coverage was 0.1, significantly less than the expected value of 0.3 (off by 67%), indicating that the majority of events were actually located closer to GT than should be expected from the model and measurement errors. The 90th percentile coverage was 0.8, close to the expected value of 1.0 (off by 20%), demonstrating

that the 90% error ellipses were indeed “honest”. However, the number of outliers (23 events with coverage larger than 2) clearly exceed the expected number at a high significance level. This indicates that the underlying “Gaussian statistics” for model and measurement error are probably inadequate for this data set. The modeling errors appear to be a conservative compromise under the condition that 90% error ellipses are “honest”. However, in order to predict “honest” 95% or 98% error ellipses the modeling errors would need to be inflated. Given the error model, 50%-60% of the time the location procedure performs better than should be expected. About 4% of the time it performs worse than should be expected.

**Table 1: Defining phases versus GT category**

<b>GT category</b>	<b># of events</b>	<b># of defining phases</b>	<b># of defining regional phases</b>	<b># of defining teleseismic phases</b>
GT0	66	11438	1054	10384
GT1	86	19640	1722	17918
GT2	131	2218	1475	743
GT5	317	54505	8963	45542
GT10	25	3116	229	2887
Total	625	90917	13443	77474

**Table 2: 90% ellipse coverages in each test using the GT0-GT10 events**

<b>test</b>	<b># of events</b>	<b>90% coverage</b>
Pn and Sn phases	571	97%
all regional phases	575	97%
all regional and teleseismic phases	625	93%

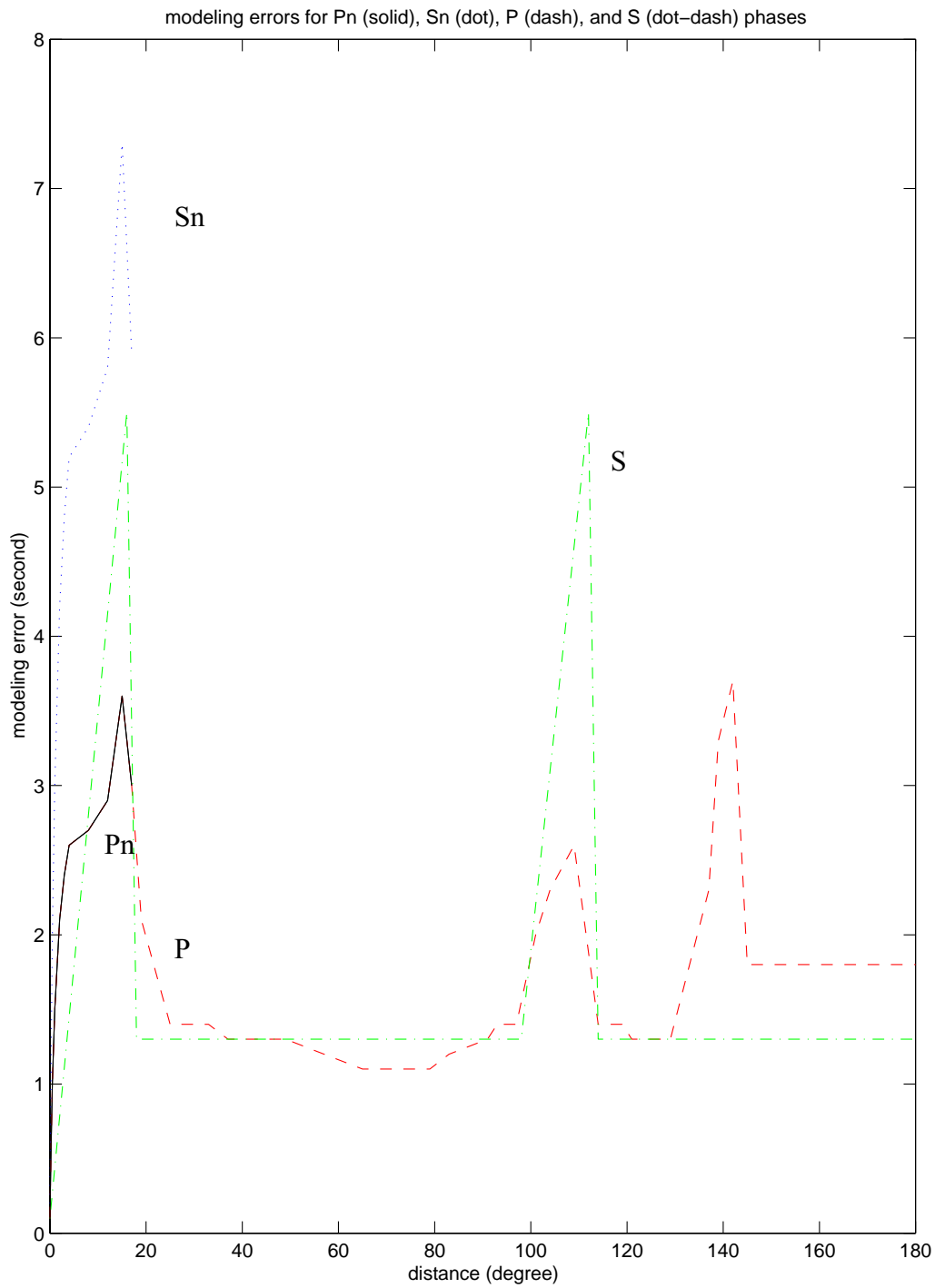


Figure 1. Modeling errors versus distance for regional and teleseismic P and S phases that are used in the current IDC system.

## 625 GT0-GT10 events in the Group-2 database

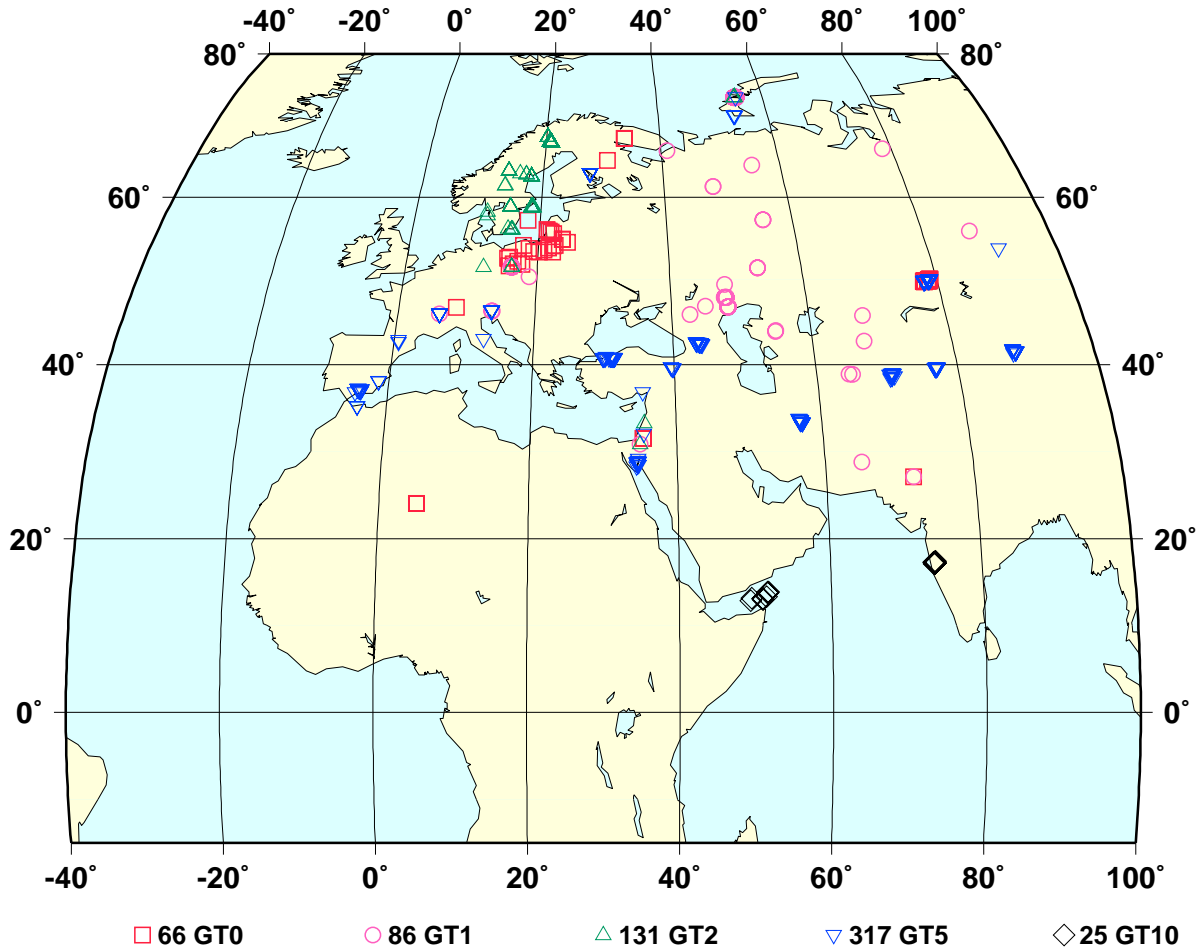


Figure 2. 625 GT0-GT10 events from the Group-2 Location Calibration Consortium Database. They are relocated using the IASPEI91 travel-time tables used in the current PIDC Operations. The 90% error ellipse coverage is 93% for these GT events, indicating the modeling error estimates used are reasonable and “honest”.

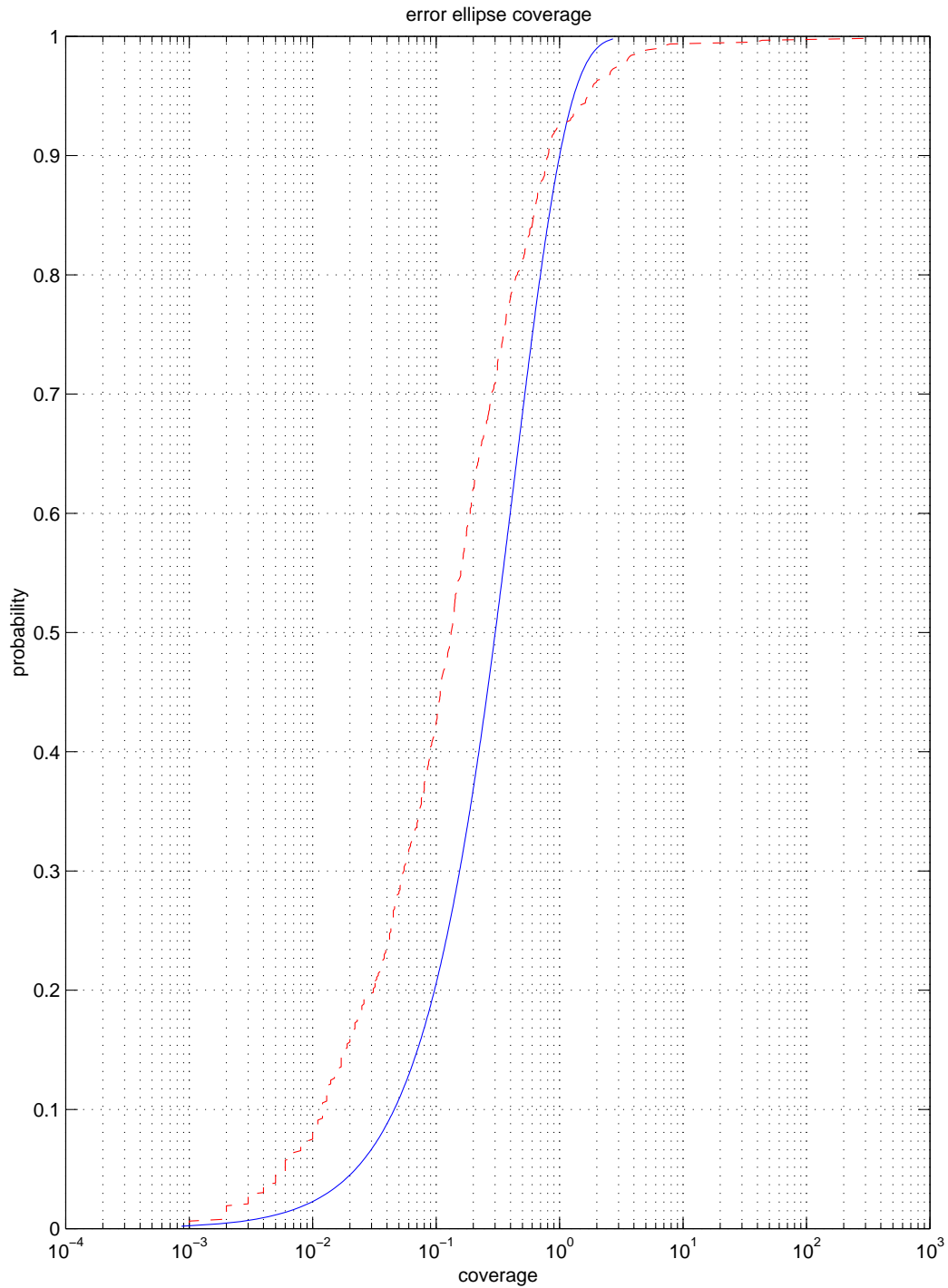


Figure 3. Error ellipse coverage for 625 events relocated using the IDC IASPEI91 travel-time table (dashed), compared to the theoretical  $\chi^2$  square distribution with 2 degrees of freedom (solid). The Kolmogorov-Smirnov Test value is 0.28, indicating fairly high risk for the assumption that the two distributions are similar. Over 50% of the events are located closer to the GT than would be expected given the model errors and uncertainties in the GT locations. 4% of the events are located much worse than would be expected.

# Memo

---

To: Keith McLaughlin  
From: István Bondár  
Date: 26 July, 2001  
Subject: *Normalizing mislocation*  
CC: Hans Israelsson, Joydeep Bhattacharyya, Xiaoping Yang

---

## Introduction

This memo describes how the 90% coverage is actually calculated by taking into account the uncertainties in reference event epicenters. Mislocations may be normalized using a similar methodology.

## Coverage parameter

The coverage parameter is calculated by substituting the reference event location into the normal equation of the ellipse:

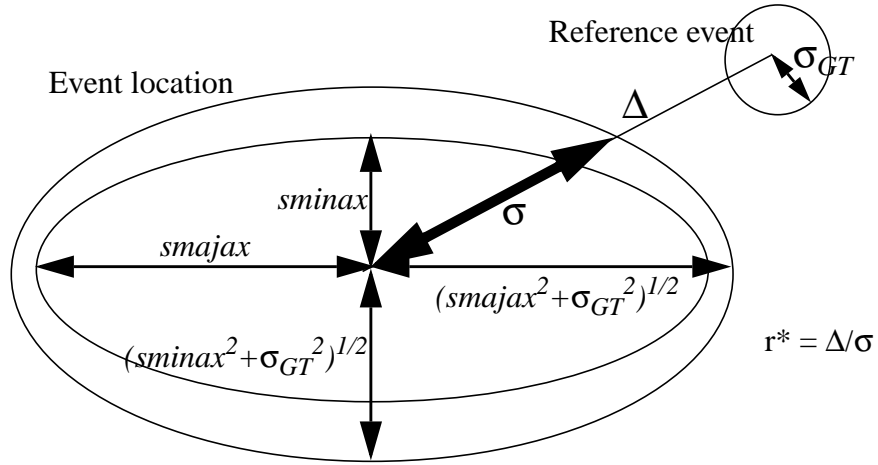
$$\vartheta = (x/smajax)^2 + (y/sminax)^2$$

where the coordinate system is centered on the event location and rotated in a way that the semi-major axis of the 90% coverage ellipse is directed to East, the semi-minor axis to North.  $x$  and  $y$  are the coordinates of the ground truth location in the above coordinate system. The coverage parameter  $\vartheta$  is less or equal than 1 if the error ellipse contains the ground truth location, and larger than 1 if the reference event falls outside the error ellipse. The coverage parameter follows a  $\chi^2$  distribution with 2 degrees of freedom.

However, the ground truth location may also have some uncertainty or limited accuracy. To take into consideration the uncertainty in the reference event location,  $smajax$  and  $sminax$  in the above equation are replaced by  $(smajax^2 + GTaccuracy^2)^{1/2}$  and  $(sminax^2 + GTaccuracy^2)^{1/2}$ .

## Mislocation normalization

Mislocation itself carries little information without accompanying error estimates. In order to present combined mislocation and uncertainty, we weight the mislocation by the extent of bias (both in the ground truth and the seismic network location) along the back-azimuth. This is illustrated in the figure below. In the coordinate system described earlier the normalized mislocation is defined as  $\Delta/\sigma$ , where  $\Delta$  is the mislocation in km and  $\sigma$  is the uncertainty along the back-azimuth in km. Therefore the normalized mislocation is a dimensionless number. It is larger than 1 if the reference event is not covered by the modified error ellipse.



Since coverage is a normalized measure of mislocation, direct comparison of coverage may be misleading; the same absolute mislocation can lead to different values of coverage if the error ellipse varies. Furthermore, if coverage is dominated by the uncertainty in the GT location, it is not possible to determine which location is better. Coverage should be used to test the underlying error model and the power of the data to test hypotheses. This modified coverage statistic assumes that GTX uncertainty,  $smajax$ , and  $sminax$  are all given at the 90'th confidence percentile.